

Gauge functions for convex cones

B. F. Svaiter*

June 20, 2011

Abstract

We analyze a class of sublinear functionals which characterize the interior and the exterior of a convex cone in a normed linear space.

keywords: normed vector spaces, cones, sublinear functionals.

Let X be a real normed linear space and K a convex closed pointed cone in X . The relation

$$x \preceq_K y \iff y - x \in K \quad (1)$$

is a partial order in X and very often one is interested in minimizing in the sense of this order a function $F : \mathbb{R}^n \rightarrow X$, that is,

$$\text{find } x \in \Omega \text{ such that } F(x) \text{ is minimal in } F(\Omega), \quad (2)$$

where Ω is a subset of \mathbb{R}^n .

Very recently Cauchy method, Newton method and Gradient Projection method were extended to problem (2) in the cases where F is differentiable, the interior of K is non-empty and $\Omega = \mathbb{R}^n$ or Ω is a convex closed set [3, 4, 1, 2]. All these methods are K -descent methods. In particular, the generated sequences will have K -smaller objective values than the initial iterate. At the heart of these extensions is a *gauge function* for K , which measures how good is a descent direction. Although the original definition considered cones in finite-dimensional spaces, its extension to an infinite dimensional setting is straightforward, and this is the case we shall discuss. Recall that the *positive polar cone* of K is K^+ ,

$$K^+ = \{x^* \in X^* \mid \langle x, x^* \rangle \geq 0, \forall x \in K\}. \quad (3)$$

*Instituto de Matemática Pura e Aplicada, (IMPA), Estrada Dona Castorina 110, Rio de Janeiro, RJ, CEP 22460-320, Brazil, benar@impa.br. The work of this author was partially supported by CNPq grants no. 474944/2010-7, 303583/2008-8 and FAPERJ grant E-26/110.821/2008.

Definition 1. A gauge function of/for a closed convex pointed cone $K \subset X$ is

$$\varphi(x) = \sup_{x^* \in C} \langle x, x^* \rangle$$

where $C \subset K^+ \setminus \{0\}$ is a weak-* compact set which generates K^+ , in the following sense

$$K^+ = \text{cl conv} (\cup_{t \geq 0} tC).$$

The rationale for the above definition is given in the next elementary result.

Proposition 2. Let $K \subset X$ be a closed convex pointed cone with a nonempty interior. Then

$$-K = \{x \in X \mid \langle x, x^* \rangle \leq 0, \forall x^* \in K^+\}, \quad (-K)^\circ = \{x \in X \mid \langle x, x^* \rangle < 0, \forall x^* \in K^+ \setminus \{0\}\},$$

where $(-K)^\circ$ stands for the interior of $-K$.

Proof. The inclusion $-K \subset \{x \mid \langle x, x^* \rangle \leq 0, \forall x^* \in K^+\}$ follows trivially from (3). To show that this inclusion holds as an equality, suppose that $y \notin -K$. Using Hahn-Banach theorem in its geometric form we conclude that there exists x^* such that

$$\langle x, x^* \rangle < \langle y, x^* \rangle, \quad \forall x \in -K.$$

As $-K$ is a cone, we conclude that $\langle x, x^* \rangle \leq 0$ for any $x \in -K$. Therefore $x^* \in K^+$ and $y \notin \{x \mid \langle x, x^* \rangle \leq 0, \forall x^* \in K^+\}$.

To prove the second equality, let

$$V = \{x \mid \langle x, x^* \rangle < 0, \forall x^* \in K^+ \setminus \{0\}\}.$$

If $x \in (-K)^\circ$ and $\langle x, x^* \rangle = 0$ for some $x^* \in K^+ \setminus \{0\}$, then there exists y in a neighborhood of x such that $y \in -K$ and $\langle y, x^* \rangle > 0$ in contradiction with the first equality of the proposition. Therefore,

$$(-K)^\circ \subset V \subset -K,$$

where the second inclusion follows from the definition of V and the first equality of the proposition. If the first inclusion is proper, there exists $z \in V \setminus (-K)^\circ$. Using again Hahn-Banach Theorem we conclude that there exists x^* such that

$$\langle x, x^* \rangle < \langle z, x^* \rangle, \quad \forall x \in (-K)^\circ.$$

Since the closure of $(-K)^\circ$ is $-K$, $x^* \leq 0$ in $-K$. Therefore $x^* \in K^+ \setminus \{0\}$ and, taking the sup on the left hand-side inequality for $x \in (-K)^\circ$ we conclude that $\langle z, x^* \rangle \geq 0$ in contradiction with the assumption $y \in V$. \square

A very natural question is: how general is the class of gauge functions? Before answering this question in Theorem 3, recall that *Fenchel-Legendre conjugate* of $f : X \rightarrow \bar{\mathbb{R}}$ is $f^* : X^* \rightarrow \bar{\mathbb{R}}$,

$$f^*(x^*) = \sup_{x \in X} \langle x, x^* \rangle - f(x)$$

and the *indicator function* of $A \subset X$ is $\delta_A : X \rightarrow \bar{\mathbb{R}}$,

$$\delta_A(x) = \begin{cases} 0, & x \in A, \\ \infty, & \text{otherwise.} \end{cases}$$

In view of the above definitions, the function φ on Definition 1 can be also expressed as

$$\varphi(x) = (\delta_C)^*(x), \quad x \in X \quad (4)$$

where we identify X with its canonical injection in to its (topological) bidual.

The aim of this note is to prove that following result:

Theorem 3. *If K is a closed convex pointed cone with a non-empty interior then φ is a gauge function of/for K if and only if it satisfies the following properties:*

1. φ is a continuous sublinear functional
2. $\varphi(x) < 0$ in the interior of $-K$;
3. $\varphi > 0$ in the complement of $-K$.

Proof. First suppose that φ is a gauge function. Then it satisfies trivially item 1. Using Proposition 2 (and the weak-* compactity of $C \subset K^+ \setminus \{0\}$) we conclude that φ satisfies item 2. Moreover, since C generate K^+ , using again Proposition 2 we conclude that

$$\varphi \geq 0 \text{ in } X \setminus (-K). \quad (5)$$

To show that φ satisfies item 3, take $x \in X \setminus (-K)$ and $x_0 \in (-K)^\circ$. Since $-K$ is closed and $x \notin -K$, there exists $\theta \in (0, 1)$ such that

$$x_\theta = (1 - \theta)x_0 + \theta x \notin -K.$$

Therefore, using (5), item 1 and item 2 and we conclude that

$$0 \leq \varphi(x_\theta) \leq (1 - \theta)\varphi(x_0) + \theta\varphi(x) < \theta\varphi(x),$$

which proves that $\varphi(x) > 0$, that is, φ satisfies item 3.

Suppose now that φ satisfies items 1, 2 and 3. Define

$$C = \{x^* \in X^* \mid x^* \leq \varphi\}, \quad \|\varphi\| = \sup_{z \neq 0} \frac{|\varphi(z)|}{\|z\|}. \quad (6)$$

Since $\varphi(x) < 0$ for some x , $0 \notin C$. Using items 1 and 2 we conclude that $\varphi \leq 0$ in $-K$. Therefore, if $x^* \in C$ then $x^* \leq 0 \in -K$ and so, $x^* \in K^+$. Altogether we have

$$C \subset K^+ \setminus \{0\}.$$

If $x^* \in C$ then $\|x^*\| = \sup_{z \neq 0} \langle x, x^* \rangle / \|z\| \leq \|\varphi\|$. Therefore, C is bounded. Since φ is sublinear

$$\varphi^*(x^*) = \begin{cases} 0, & x^* \in C \\ \infty, & \text{otherwise.} \end{cases}$$

As φ^* is convex and weak-* lower semicontinuous, C is convex and weak-* closed. Hence C is weak-* compact. Moreover, by Fenchel-Moreau Theorem, for any $x \in X$

$$\varphi(x) = \varphi^{**}(x) = \sup_{x^* \in C} \langle x, x^* \rangle = \max_{x^* \in C} \langle x, x^* \rangle \quad (7)$$

where the second equality follows from the above expressions for φ and the third one from the weak-* compactness of C .

Let V be the cone generated by C ,

$$V = \text{conv}(\cap_{t \geq 0} tC) = \cap_{t \geq 0} tC$$

where the second equality follows from the convexity of C . As $C \subset K^+$, $V \subset K^+$. If $V \neq K^+$, there exists $w_0^* \in K^+ \setminus V$ and

$$\{tw_0^* \mid t \geq 0\} \cap C = \emptyset.$$

Using the convexity and weak-* compactness of C we conclude that there exists x_0 and b such that

$$\langle x_0, z^* \rangle < b \leq \langle x_0, tw_0^* \rangle, \quad \forall t \geq 0, z^* \in C.$$

Therefore $b \leq 0$, $\langle x_0, w_0^* \rangle \geq 0$ and, in view of (7), $\varphi(x_0) < 0$. Hence, by item 2, $x_0 \in (-K)^\circ$. However, $w_0^* \in K^+ \setminus \{0\}$ and $\langle x_0, w_0^* \rangle \geq 0$, in contradiction with Proposition 2. So, the assumption $K^+ \neq V$ is false and K^+ is the cone generated by C , which completes the proof that φ is a gauge function with C given by (6). \square

In some cases, given a gauge function is easy to provide alternative choices for C (not necessarily weak-* compacts), as shown in the next example, Hiriart-Urruty's *oriented distance function*, introduced in [5] in the framework of nonsmooth scalar optimization in Banach spaces.

Proposition 4. *Suppose that K is a pointed convex closed cone with a non-empty interior, and let $\varphi : X \rightarrow \mathbb{R}$ be given by*

$$\varphi(x) = d(x, -K) - d(x, X \setminus -K)$$

Then

$$\varphi(x) = \sup_{x^* \in C} \langle x, x^* \rangle$$

for $C = \{x^ \in K^+ \mid \|x^*\| = 1\}$.*

Proof. Take $x \notin -K$, and let $r = d(x, -K)$. Define

$$A = -K + B(0, r).$$

Note that A is an open convex set, $0 \in A$ and $x \in \bar{A}$. Let g be the Minkowski functional of A ,

$$g(z) = \inf\{t > 0 \mid t^{-1}z \in A\}, \quad z \in X.$$

Using Hahn-Banach Theorem we conclude that there exists $x^* \in X^*$ such that $x^* \leq g$, $x^*(x) = g(x) = 1$. Since $-K \subset A$, $x^* \in K^+$. Since $B(0, r) \subset A$, $\|x^*\| \leq 1/r$. As $x^*(x) = 1$, there exists a sequence $\{y_n\}$ in $-K$ such that

$$\|x - y_n\| \rightarrow r, \quad \text{as } n \rightarrow \infty$$

Hence

$$\langle x - y_n, x^* \rangle = 1 - \langle y_n, x^* \rangle \geq 1.$$

Combining the two above equations we conclude that $\|x^*\| \geq 1/r$. Therefore, $\|x^*\| = 1/r$,

$$rx^* \in C, \quad \langle rx^*, x \rangle = r$$

which proves that

$$\varphi(x) = r \leq \sup_{y^* \in C} \langle x, y^* \rangle.$$

To prove the converse, using the definition of C and the sequence $\{y_n\}$ we have

$$\sup_{y \in C} \langle x, y^* \rangle = \sup_{y \in C} \langle x - y_n, y^* \rangle + \langle y_n, y^* \rangle \leq \|x - y_n\|$$

so $\varphi(x) = \sup_{y^* \in C} \langle x, y^* \rangle$.

Take $x \in (-K)^\circ$ and let $\rho = d(x, X \setminus -K)$. There exists a sequence $\{z_n\}$ in $X \setminus -K$ such that

$$\|x - z_n\| \rightarrow \rho, \quad \text{as } n \rightarrow \infty$$

As $z_n \notin -K$, for each n there exists x_n^* such that $\|x_n^*\| = 1$ and

$$\langle y, x_n^* \rangle < \langle z_n, x_n^* \rangle, \quad \forall y \in -K.$$

The above inequality trivially implies $x_n^* \in K^+$, $\langle z_n, x_n^* \rangle > 0$ and so

$$\langle x, x_n^* \rangle = \langle x - z_n, x_n^* \rangle + \langle z_n, x_n^* \rangle \geq -\|x - z_n\|.$$

Hence

$$\sup_{y^* \in C} \langle x, y^* \rangle \geq -\rho = \varphi(x).$$

To prove the converse, note that for any $x^* \in C$, $\|x^*\| = 1$,

$$\langle x, x^* \rangle + \rho = \sup_{\|y\| < \rho} \langle x + y, x^* \rangle \leq \sup_{y' \in -K} \langle y', x^* \rangle = 0$$

which trivially implies the desired inequality. \square

References

- [1] L. M. Graña Drummond and A. N. Iusem. A projected gradient method for vector optimization problems. *Comput. Optim. Appl.*, 28(1):5–29, 2004.
- [2] J. Fliege, L. M. Graña Drummond, and B. F. Svaiter. Newton’s method for multiobjective optimization. *SIAM J. Optim.*, 20(2):602–626, 2009.
- [3] Jörg Fliege and Benar Fux Svaiter. Steepest descent methods for multi-criteria optimization. *Math. Methods Oper. Res.*, 51(3):479–494, 2000.
- [4] L. M. Graña Drummond and B. F. Svaiter. A steepest descent method for vector optimization. *J. Comput. Appl. Math.*, 175(2):395–414, 2005.
- [5] J.-B. Hiriart-Urruty. Tangent cones, generalized gradients and mathematical programming in Banach spaces. *Math. Oper. Res.*, 4(1):79–97, 1979.